



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12Q 1/68		A2	(11) International Publication Number: WO 98/58083
			(43) International Publication Date: 23 December 1998 (23.12.98)
<p>(21) International Application Number: PCT/US98/12420</p> <p>(22) International Filing Date: 15 June 1998 (15.06.98)</p> <p>(30) Priority Data: 08/876,766 16 June 1997 (16.06.97) US</p> <p>(71) Applicant (for all designated States except US): MEDICAL COLLEGE OF OHIO [US/US]; 3000 Arlington Avenue, P.O. Box 10008, Toledo, OH 43699-0008 (US).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (for US only): WILLEY, James, C. [-US]; 4235 Deepwood Lane, Toledo, OH 43614 (US). CRAWFORD, Erin, L. [-US]; 1067 Sandusky Place #5, Perrysburg, OH 43551 (US). DEMUTH, Jeffery, P. [-US]; 3129 Sherbrooke Road, Toledo, OH 43606 (US). JACKSON, Clara, M. [-US]; 1148 Hugo Street, Maumee, OH 43537 (US). WEAVER, David, A. [-US]; 10702 Avenue Road, Perrysburg, OH 43551 (US).</p> <p>(74) Agents: MARTINEAU, Catherine, B. et al.; Emch, Schaffer, Schaub & Porcello Co., P.O. Box 916, Toledo, OH 43697-0916 (US).</p>		<p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p>Published Without international search report and to be republished upon receipt of that report.</p>	
<p>(54) Title: METHOD FOR QUANTITATIVE MEASUREMENT OF GENE EXPRESSION USING MULTIPLEX COMPETITIVE REVERSE TRANSCRIPTASE-POLYMERASE CHAIN REACTION</p> <p>(57) Abstract</p> <p>A method and apparatus for quantitative measurement of gene expression through multiplex competitive reverse transcriptase polymerase chain reaction amplification are disclosed. The method and apparatus are especially useful for analysis of small specimens of cells and tissues.</p>			
<pre> graph TD A[Extraction of RNA] --> B[Reverse Transcription] B --> C[PCR Amplification (Native cDNA and Mutated Standards)] C --> D[Enzyme Restriction] D --> E[Agarose Gel Electrophoresis] E --> F[Densitometry on Negative Photograph of Gel] F --> G[Ratio of Target Gene to "Housekeeping" Gene] </pre>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
C	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

DESCRIPTION

METHOD FOR QUANTITATIVE MEASUREMENT
OF GENE EXPRESSION USING MULTIPLEX COMPETITIVE
REVERSE TRANSCRIPTASE-POLYMERASE CHAIN REACTION

5

TECHNICAL BACKGROUND

This is a continuation in part of Serial. No. 08/043,390 filed April 6, 1993 and a continuation-in-part of Serial No. 08/118,434 filed January 28, 1994.

10 The present invention was made under research grant number ES02679 and ES01247 from the National Institute of Health, Grant No. RR00044 from the Division of Research Resources, Health Institute Contract 91-2 and International Lead Zinc Organization Contract CH611 who may have certain rights thereto. The present invention relates 15 generally to a method for the quantitative measurement of cellular levels of RNA following reverse transcription and polymerase chain reaction (PCR) amplification.

The PCR techniques are generally described in U.S. Patent Nos. 4,683,195; 4,683,202; and 4,965,188. The PCR technique generally 20 involves a process for amplifying any desired specific nucleic acid sequence contained within a nucleic acid sequence which is within a nucleic acid molecule. The PCR process includes treating separate complementary strains of the nucleic acid with an excess of two oligonucleotide primers. The primers are extended to form 25 complementary primer extension products which act as templates for synthesizing the desired nucleic acid sequence. The PCR process is carried out in a simultaneous step-wise fashion and can be repeated as often as desired in order to achieve increased levels of amplification of the desired nucleic acid sequence. According to the PCR process, the 30 sequence of DNA between the primers on the respective DNA strains are

amplified selectively over the remaining portions of the DNA and selected sample. The PCR process provides for the specific amplification of a desired region of DNA.

The yield of product from PCR increases exponentially for an indefinite number of cycles. At some point and for uncertain reasons, the reaction becomes limited and PCR product increases at an unknown rate. Consequently, the yield of amplified product has been reported to vary by as much as 6-fold between identical samples run simultaneously. (Gilliland, G., et al., *Proc. Natl. Acad. Sci.* 87:2725-2729, 1990). (These publications and other reference materials have been included to provide additional details on the background of the invention and, in particular instances, the practice of the invention, and all are expressly incorporated herein by reference.) Therefore, after a certain number of PCR cycles, the initial concentrations of target DNA cannot be accurately determined by extrapolation. In an attempt to make PCR quantitative, various investigators have analyzed samples amplified for a number of cycles known to provide exponential amplification (Horikoshi, T., et al., *Cancer Res.* 52:108-116 (1992); Noonan, K.E., et al., *Proc. Natl. Acad. Sci.* 87:7160-7164 (1990); Murphy, L.D., et al., *Biochemistry* 29:10351-10356 (1990); Carre, P.C., et al., *J. Clin. Invest.* 88:1802-1810 (1991); Chelly, J., et al., *Eur. J. Biochem* 187:691-698 (1990); Abbs, S., et al., *J. Med. Genet.* 29:191-196 (1992); Feldman, A.M. et al., *Circulation* 83:1866-1872 (1991). In general, these analyses are done early in the PCR process when the PCR product is measurable by use of radiolabeled probes and autoradiography but not by spectrophotometry or densitometry of ethidium bromide stained gels. The use of radioactivity is inconvenient, expensive, and presents safety concerns. Also, the exponential phase must be defined for each set of experimental conditions, requiring additional cost in time and materials.

Another development is competitive PCR, wherein PCR is conducted in the presence of single base mutated competitive templates (*Gilliland, supra*; Becker-Andre, et al., *Nucleic Acids Res.* 17:9437-9446 (1989)). A known amount of competitive template is co-amplified with 5 an unknown amount of target sequence. The competitor is the same sequence (except for single base mutation or deletion of a portion of the sequence) as the target, uses the same primers for amplification as the target cDNA, and amplifies with the same efficiency as the target cDNA. The starting ratio of target/standard is preserved throughout the entire 10 amplification process, even after the exponential phase is complete.

Competitive PCR is discussed in general in Siebert, P.D., et al., *Nature* 359:557-558 (1992); Siebert, P.D., et al., *Bio Techniques* 14:244-249 (1993), and Clontech Brochure, 1993, Reverse Transcriptase-PCR (RT-PCR). However, competitive PCR alone does not adequately control 15 for variation in starting amounts of template. Degradation of samples and pipetting errors can lead to variation. When using Northern analysis to measure gene expression, it is possible to overcome these problem by probing the same blot for both a target gene and a "housekeeping" gene which is not expected to vary among tissue 20 samples or in response to stimuli. The "housekeeping" gene acts as a denominator in determining the relative expression of a target gene. In attempts to apply this concept, other investigators have PCR-amplified in separate tubes. However, when the two genes are amplified in separate tubes, intertube variation in amplification conditions and 25 pipetting errors are unavoidable. While non-competitive multiplex PCR, where the target and "housekeeping" gene are amplified in the same tube, has also been described in *Noonan, supra*, this method is inconvenient because it requires the generation of standard curves to determine the exponential range of amplification nuclides.

Therefore, there is a need for quantitative measurement of gene expression technique which has none of the above-described drawbacks and which can be performed by a technician with standard training. The present invention addresses these needs in the art by providing a 5 technique which can be utilized with any PCR process and which can be performed in a simple and straightforward manner. The present invention involves a dramatic improvement over previously described approaches to DNA analysis and the PCR techniques.

10

SUMMARY OF THE INVENTION

The method of the present invention is an improvement upon the PCR amplification process that allows simultaneous amplification of a "target gene", a "housekeeping" gene and competitive templates for each of these genes. According to the present invention, the terms 15 "target DNA sequence" and "target gene" generally refer to a gene of interest for which there is a desire to selectively amplify that gene or DNA sequence. The term "housekeeping" gene refers to genes that are suitable as internal standards for amount of RNA per PCR reaction. In a general and overall sense, a key to the present invention is the 20 simultaneous use of primers for a target gene, primers for a housekeeping gene, and two internal standard competitive templates comprising mutants of the target gene and housekeeping gene. These mutations can be point mutations, insertions, deletions or the like.

In a broad sense, the present invention is directed to a method for 25 quantifying the amount of a target DNA sequence within an identified region of a selected cDNA molecule that is present within a heterogenous mixture of cDNA molecules. It is to be understood that more than one targeted gene and/or housekeeping gene can be utilized and further that quantitation of such additional target and/or housekeeping genes will 30 necessitate the further inclusion of an internal standard competitive

template comprising a mutation of that additional target and/or housekeeping gene. It is to be understood that the mutated competitive templates comprise at least one nucleotide that is mutated relative to the corresponding nucleotide of the target sequence. It is to be noted that

5 mutation of only a single nucleotide that is complementary to the corresponding nucleotide of the housekeeping gene sequence is required for the successful practice of the present invention. However, it is understood that longer deletions, insertions or alterations are useful in the present invention. The target gene primers (which serve as primers for

10 both the native and competitive templates of the target gene), housekeeping gene primers (which serve as primers for both the native and competitive template of the housekeeping gene), competitive template of the target gene, and competitive template of the housekeeping gene are subjected to a PCR process along with native

15 cDNA which contains the DNA for both the target gene and the housekeeping gene. The PCR process provides cDNA products of 1) native cDNA of the target gene and the housekeeping gene and 2) mutated competitive template cDNA of the target gene and the housekeeping gene. The cDNA products are isolated using methods

20 suitable for isolating cDNA products. The relative presence of the native cDNA products and the mutated cDNA products are detected by measuring the amounts of native cDNA coding for the target gene and mutated cDNA coding for the competitive template of the target gene as compared to the amounts of native cDNA coding for the housekeeping

25 gene and mutated cDNA coding for competitive template of the housekeeping gene.

In a further aspect, the present invention concerns novel nucleic acid primer molecules for use in PCR processes for determining the amount of a target DNA sequence.

According to the present invention herein "a sample" generally indicates a sample of tissue or fluid isolated from a plant, individual or *in vitro* cell culture constituents.

The terms primers, nucleic acids and oligonucleotides are understood to refer to polyribonucleotides and polydeoxyribonucleotides and there is no intended distinction in the length of sequences referred to by these terms. Rather, these terms refer to the primary structure of the molecule. These terms include double and single stranded RNA and double and single stranded DNA. It is to be understood that the oligonucleotides can be derived from any existing or natural sequence and generated in any manner. It is further understood that the oligonucleotides can be generated from chemical synthesis, reverse transcription, DNA replication and a combination of these generating methods. The term "primer" generally refers to an oligonucleotide capable of acting as a point of initiation of synthesis along a complementary strand when conditions are suitable for synthesis of a primer extension product. The synthesizing conditions include the presence of four different deoxyribonucleotide triphosphates and at least one polymerization-inducing agent such as reverse transcriptase or DNA polymerase. These are present in a suitable buffer which may include constituents which are co-factors or which affect conditions such as pH and the like at various suitable temperatures. It is understood that while a primer is preferably a single strand sequence, such that amplification efficiency is optimized, other double stranded sequences can be practiced with the present invention.

The terms "target gene", "sequence" or "target nucleic acid sequence" are meant to refer to a region of a oligonucleotide which is either to be amplified and/or detected. It is to be understood that the target sequence resides between the primer sequences used in the amplification process.

According to one embodiment of the present invention, quantitative gene expression is measured by multiplex competitive PCR amplification of a) cDNA from at least one target gene of interest and at least one "housekeeping" gene and b) internal mutated standard competitive templates comprising base mutants of the target gene of interest and the "housekeeping" gene cDNA that cause either a loss or gain of a restriction endonuclease recognition site. According to another embodiment, the method comprises the PCR amplification of a) cDNA from at least one target gene of interest and at least one "housekeeping" gene and b) competitive templates comprising sequences of the target gene of interest and the "housekeeping" gene that have been artificially shortened. These shortened sequences retain sequences homologous to both the target gene and the housekeeping gene primers used in PCR amplification. RNA extracted from sample cells or tissues are reverse transcribed. Serial dilutions of cDNA are PCR amplified in the presence of oligonucleotides homologous to the target gene and the "housekeeping" gene, and quantified amounts of internal mutated standard competitive templates. The amplified DNA is restriction digested and electrophoresed on an agarose gel stained with ethidium bromide, separating native from mutated products. Densitometry is performed to quantify the bands. This technique to measure the relative expression of a target gene to a "housekeeping" gene is precise and reproducible for studies done with the same master mixture and dilution of internal standards. Ratios of relative gene expression vary less than 25% from the mean. This technique is useful to measure changes in gene expression. This method is particularly useful when the amount of study sample is limited or the level of gene expression is low.

According to another embodiment of the present invention, the quantity of gene expression is measured by conducting, prior to the simultaneous polymerase chain reaction amplification for each mutated

competitive template, the following steps: two initial polymerase chain reactions conducted using an outer primer and inner single base mismatched mutated internal standard competitive template primer to produce two overlapping DNA segments; isolation and purification of the

5 overlapping DNA fragments of the initial polymerase chain reaction; using a polymerase chain reaction to amplify each of the two overlapping DNA fragments using the outer primers only; conducting a first polymerase chain reaction amplification without the primers to allow for heterodimer formation; and purifying and amplifying the polymerase chain reaction

10 products of the previous step and thereafter diluting to use as competitive templates.

According to another embodiment of the present invention the competitive templates for many different genes are included in the same competitive template mixture, rather than the competitive templates for

15 just the housekeeping gene and a single target gene. Also, Quantitative RT-PCR may be done successfully without amplifying the housekeeping gene and target gene in the same tube following the methods described herein. The present invention further relates to an automated molecular phenotyping apparatus.

20 It is, therefore, the object of the present invention to provide an improved method for quantitative measurement of gene expression.

It is a further object of the present invention to provide a method for quantitative PCR-based measurement of gene expression that is suitable as a commercial process.

25 These and other objects, features and many of the attendant advantages of the invention will be better understood upon a reading of the following detailed description when considered in connection with the accompanying drawings wherein.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a schematic diagram of the work scheme for PCR quantification of relative target gene expression.

Fig. 2 is a photograph showing serial dilutions of BEP2D human papillomavirus-immortalized human bronchial epithelial cell line cDNA (representing 0.25 μ g to 0.05 μ g total RNA) that were co-amplified with constant amounts of each single base mutated internal standard (10 attamoles each), subjected to EcoRI restriction endonuclease digestion and electrophoresed on an agarose gel. A negative photograph of this gel was then subjected to densitometry in order to quantify bands.

Fig. 3 is a graph showing the native product (□GSH-Px (glutathione peroxidase) and ■GAPDH (glyceraldehyde-3-phosphate dehydrogenase)) vs. total RNA. With small amounts of starting RNA (first three squares for both GSH-Px) and GAPDH), the reactions are exponential throughout, however, with increasing amounts of RNA, the reaction becomes non-exponential at some point during amplification, resulting in less product formation than expected in a non-limited reaction. The straight lines represent the theoretical amounts of PCR product (either GSH-Px or GAPDH) which would have formed if amplification remained exponential throughout the amplification process.

Fig. 4 is a graph showing Native/Mutated amplification products (□GSH-Px and ■GAPDH) vs. Total RNA. The relationship remains linear for both GSH-Px ($r^2 = .982$) and GAPDH ($r^2 = .973$) throughout the range of total RNA studied. This is consistent with the findings of Gilliland et al., *Proc. Natl. Acad. Sci.*, 87:2725-2729, 1990. This relationship attests to the quantitative nature of multiplex competitive PCR.

Fig. 5 is a photograph showing a) Northern analysis of RNA obtained from BEP2D cells that were treated with 0.1% DMSO as a control, or β -naphthoflavone in an effort to induce cytochrome p450 IA1 (CYP1A1); and b) DNA PCR-amplified from serial dilutions of cDNA from

the same cells as in a. The cDNA was co-amplified in the presence of competitive template internal standards for GAPDH and CYPIAI and oligonucleotide primers for these genes.

Fig. 6 is a schematic illustration of an apparatus for quantitatively
5 measuring gene expression.

DETAILED DESCRIPTION OF THE INVENTION

For many years, gene expression has been measured through quantification of RNA by Northern or dot blot analysis. These techniques require the amount of RNA obtainable from at least 10^5 cells for each
10 measurement. Often, a biopsy will provide only the number of cells necessary for a histological diagnosis and this is often far less than 10^5 cells. Recently developed PCR techniques allow measurement of RNA levels in as few as 100 cells. However, techniques described thus far allow only qualitative, not quantitative measurement.

15 The present invention relates to a method using multiplex competitive reverse-transcriptase polymerase chain reaction amplification to simplify and improve quantitative measurement of gene expression. DNA extracted from samples is reverse transcribed and then subjected to PCR amplification in the presence of primers for both a
20 "housekeeping" gene and a target gene of interest.

25 The expression of genes is measured by comparing the amount of at least one target gene to that of a "housekeeping" gene. There are various suitable housekeeping genes which are useful with the present invention, including, for example, such housekeeping genes that have been used as internal standards in Northern analyses of gene expression, including GAPDH, β -actin, 28S RNA and 18S RNA and ribonuclear proteins (Devereux et al., Nucleic Acids Res. 12:387 (1984); Barbu et al., Nucleic Acids Res. 17:7115 (1989). According to one embodiment of the present invention, synthesized oligonucleotides homologous to any
30 sequences containing a known restriction endonuclease recognition site

or any sequence containing one or two-basepair mismatch for a known restriction endonuclease site that is present in the housekeeping gene can be utilized. The application of these restriction endonuclease recognition sites is to either mutate the naturally occurring sites to non-
5 recognition sites or to mutate the mismatch sites to match sites, in either case creating mutant sequences suitable for internal mutated standard competitive templates. The particular sites in the housekeeping gene used for analysis of any particular other gene depends on the match and mismatch sites that are present in the other gene. One determinant is
10 the size of the DNA fragments that are generated from the housekeeping gene and the target gene. It is desired that these fragments separate well on gel electrophoresis.

Further, all oligonucleotides that contain sequences homologous to sequences in the genes for the housekeeping genes can be used in the
15 present invention. Such homologous sequences may be used to generate artificially shortened competitive templates to the housekeeping genes generated according to the method described by Celi et al., Nucleic Acids Res. 21:1047 (1993).

To identify and match one or two base mismatch sequences for all
20 known recognition sites, it is possible to use the Map program within the Genetics Computer Group software package (Devereux et al., *supra*, 1984). The cDNA sequences are obtained for each gene, then each gene is evaluated for the presence of the match of one or two base pair mismatch sequences for every known restriction endonuclease.
25 According to the present invention, it is possible to use every gene containing any of these recognition sequences or one or two base pair mismatches of these sequences.

The present invention of multiplex competitive PCR (for any given master mixture and dilution of mutated standards) yields reproducible
30 ratios of target gene/housekeeping gene. Because the relationship

between the amount of native product vs. total starting RNA does not remain linear with increasing amounts of RNA (Fig. 3), it is clear that the amplification is not exponential throughout. However, consistent target/housekeeping gene ratios are obtained with the use of a 5 competitive template for each. The quantitative results are illustrated by the linear relationship between the native/mutated template ratio vs. total starting RNA for each of the genes (Fig. 4). This illustrates the utility of competitive template internal controls; it is not necessary to remain in the exponential phase of amplification to quantify relative amounts of gene 10 expression. The reproducibility of the ratio between samples in the same study allows for the use of fewer dilution tubes. Only one tube, in which all bands are quantifiable, is necessary for measurement of gene expression. This simplifies the procedure and permits the evaluation of many different samples at one time.

15 According to one embodiment of the present invention there may be a choice of restriction endonuclease recognition sites used to separate the mutated from native products. In addition, the ultimate length of the PCR product present after restriction endonuclease digestion is a factor to consider. In certain embodiments, agarose gels are more preferable 20 to use than polyacrylamide gels, but require greater DNA fragment size differences for adequate separation (approximately 50-100 base pair differences). The method of the present invention can be further simplified by using the same recognition site for both the target and housekeeping gene. EcoRI sites are often appropriate. In a specific 25 embodiment of the present invention, GAPDH competitive templates were prepared that separate from native GAPDH on the basis of EcoRI or BamHI digestion. However, it is to be understood that separation on the basis of EcoRI or other restriction endonuclease digestion that is compatible with a greater number of genes may also occur.

A large variation (up to 200-fold) in the target housekeeping (GSH-Px/GAPDH) gene expression ratio has been observed in studies performed on BEP2D cDNA using different master mixtures and dilutions of mutated standards. Since samples that underwent reverse transcription separately gave similar results, the amount of variability introduced at this step is small. Any differences in reaction conditions will equally affect amplification of the competitive template as well as the native template and thus the ratio between the two will remain constant. Therefore, the variability likely results from differences in the amount of internal mutated standard competitive template in the reaction. The concentration of competitive template is so small (femtomolar range) that any change in the number of molecules present in the reaction would introduce a large source of error.

The method of the present invention is precise in any given study as illustrated by the reproducibility between samples using the same master mixture with mutated internal standard competitive templates. Therefore, according to a preferred embodiment, it is desirable that comparative samples be run simultaneously using the same master mixture with the same dilution of internal mutated standard competitive templates.

In a preferred embodiment, the reaction mixture includes the following: a) competitive template for a housekeeping gene, b) competitive template for a target gene of interest, c) oligonucleotide primers for PCR amplification of the housekeeping gene, d) oligonucleotide primers for PCR amplification of the target gene of interest, e) reaction buffer, and f) cDNA from the test sample.

It is also preferred that the PCR conditions be standardized for each experiment by using, for example, a master mixture containing 1x PCR buffer (for example, 50 mM KCl, 10 mM Tris-HCl, 1.5 mM MgCl₂), primers coding for the target gene and the housekeeping gene (25

pmoles of each), 0.2 mM dNTP's (A,T,C,G), and constant amounts of each competitive template per 100 μ L reaction mixture. TaqDNA polymerase (2.5 units) is added to each 100 μ L reaction prior to amplification. PCR amplification is preferably carried out for 35-38 cycles 5 at 94°C for one min., 60°C for one min., and 72°C for one min., or as is optimal for the particular region amplified. After amplification, PCR products are preferably heated for 10 min. and then cooled slowly in order to maximize heterodimer formation.

For reaction using competitive templates mutated according to 10 Higuchi, samples (40 μ L) from each PCR tube are restriction endonuclease digested for 12-16 hours. These products are electrophoresed on a 3% Nusieve, 1% LE agarose ethidium bromide stained gel for 2-3 hrs. at 60V. A negative photograph is taken of the gel using Polaroid 665 positive/negative instant film.

15 The negative photograph is subjected to densitometry using, for example, Zeineh Soft Laser Scanning Densitometer Model SLR 2D/1D using Zeineh 1D Autostepover Videophoresis Program Software (Biomed Instruments, Fullerton, California). Alternatively, the stained gel is evaluated densitometrically directly using a digital camera, or evaluated 20 on an automated sequencing gel (such as that offered by Applied Biosystems, Inc.). Areas under each curve are calculated and used for quantification. Corrections are made for relative band sizes and heterodimer formation. Results are expressed as target gene to housekeeping gene relative ratios.

25 Multiplex competitive PCR improves and simplifies quantitation of gene expression. Gene expression can be quantitated in very small samples of tissue or cells without resorting to radiolabeling. As a result, multiplex reverse transcription PCR is less expensive and safer to use than radiolabeling. The results are reproducible for examples using the

same master mixture and dilutions of internal mutated standard competitive templates.

It is to be understood that according to the method of the present invention, all oligonucleotides homologous to each strand of the cDNA of known or potential housekeeping genes (including but not restricted to the human, mouse and rat GAPDH, β -actin, 28S RNA, 18S RNA, and all ribonuclear protein genes) and containing restriction endonuclease recognition site sequences or one or two base pair mismatches for restriction endonuclease recognition sequences are useful in the practice 5 of the present invention. The oligonucleotides are used to prepare competitive templates of housekeeping genes for use in quantitative PCR.

It is to be further understood that according to the method of the present invention, all oligonucleotides that contain sequences homologous to sequences in known or potential housekeeping genes 10 (including but not restricted to GAPDH, β -actin, 28S RNA, 18S RNA, and all ribonuclear protein genes) are useful in generating artificially shortened competitive templates. The oligonucleotides are used to prepare competitive templates of housekeeping genes for use in the present invention.

20 It is contemplated that uses of this inventive technique include:
a) evaluating gene expression from tissues obtained by endoscopic biopsy (brush or forceps), needle aspiration, and bone marrow biopsy; b) quantification of reporter gene expression in transient transfection assays; and c) quantification of transfected genes following gene 25 therapy.

While a method described in Morales, M.J., and Gottlieb, D.I., (A polymerase chain reaction-based method for detection and quantification of reporter gene expression in transient transfection assays, Analytical Biochemistry, 210, 188-194 (1993)), allows measurement of 30 transcription from the reporter gene, it does not control for variation

resulting from differences in amount of RNA included in the assay. The method of the present invention is further useful to evaluate simultaneously both the exogenous reporter gene and an endogenous housekeeping gene, such as GAPDH RNA from the transfected cell.

- 5 Although numerous different mutations in the cystic fibrosis transmembrane conductance regulator gene (CFTR) have been reported to be associated with disease, the most common disease-associated mutation is a 3 base deletion at position 508. It is possible to prepare oligonucleotide primers that result in amplification of only the abnormal
- 10 508 deleted gene or only the normal CFTR gene using recently described methods (Cha, R.S., Zarbl, H., Keohavong, P., Thilly, W.G., match amplification mutation assay (MAMA): application to the c-Ha ras gene, PCR methods and applications, 2:14-20 (1992). It is further contemplated that it is possible to measure the relative amount of
- 15 exogenous CFTR gene per cell using the method of the present invention. Without comparing the amount of abnormal CFTR to some internal standard, it would not be possible to quantify the efficiency of transfection. The samples are too small to measure RNA by Northern analysis.
- 20 The method of the present invention is useful to develop PCR systems for selective amplification of exogenous normal dystrophin gene in the presence of mutated endogenous gene. In the case of dystrophin, the disease results from relatively large deletions. Several different ones have been described. Using the method of the present invention,
- 25 different PCR systems may be developed for selective amplification of the transfected normal gene and constitutive abnormal gene for each of these cases. This may be done by choosing primers that span the deleted region.

It should be further understood that according to the method of the present invention, more than one gene can be evaluated at the same time.

METHODS

5 Purified deoxyribonucleotides obtained from Pharmacia (Piscataway, N.J.) were diluted to a stock solution of 10 mM. Recombinant *Thermus aquaticus* DNA polymerase (Taq polymerase), Avian myeloblastosis virus (AMV) reverse transcriptase, and ribonuclease inhibitor (RNasin) were obtained from Promega (Madison, WI.). EcoRI 10 enzyme was obtained from USB (Cleveland, Ohio). Primers were prepared on an Applied Biosystems model 391 PCR-Mate EP TM synthesizer. PCR was performed in a Perkins, Elmer, Cetus DNA Thermal Cycler 480. The other buffers and solutions used were from various sources and were molecular biology grade.

15 Studies were performed on a human papillomavirus-immortalized human bronchial epithelial cell line (BEP2D) (Willey et al, *Cancer Res.* 5 1:5370-5377, 1990).

20 The isolation of RNA was as follows: RNA was isolated based on the method described by Chomczynski and Sacchi (*Analytical Biochemistry* 16 2:156-159, 1987) Culture medium was removed from flasks containing the BEP2D cell line. Immediately GIT (4.0 M guanidinium thiocyanate, 0.1 M Tris Cl Ph= 7.5, 1% beta-mercaptoethanol) buffer was placed on the cells (approximately 500 μ L per 5-10 million BEP2D cells). Each 500 μ L of GIT buffer containing the 25 lysed cells was then transferred to a 1.5 mL microfuge tube. To each microfuge tube was added sequentially 50 μ L of 2M Na acetate pH= 4, 500 mL of water saturated phenol and 100 mL of chloroform-isoamyl alcohol mixture (49:1). The tubes then were shaken thoroughly, placed on ice for 15 min, and microcentrifuged for 20 min at 14,000 RPM and 30 4°C. The aqueous phase of each tube was transferred to a fresh tube

and the above extraction was repeated. Again, the aqueous phase of each tube was transferred to a fresh tube with isopropanol (500 μ L), and placed at -70°C for 15 min. The tubes were then microcentrifuged for 20 min at 14,000 RPM and 4°C. The RNA was washed twice with 70% 5 ethanol and vacuum dried. RNA was taken up in 0.1% diethyl pyrocarbonate (DEPC) treated H₂O and quantified by spectrophotometry (Gilford Instrument Spectrophotometer 260).

The reverse transcription was conducted as follows: the extracted RNA was placed in a sterile microfuge tube. For each 1 μ g of RNA, 0.2 10 mg oligo dT was added. This was heated at 65°C for 5 min and placed on ice for one min. To this was added 2 μ L 1- mM dNTP's, 2 μ L reverse transcriptase (RT) buffer (50.0 mM Tris, 40.0 mM KCl, and 80 mM MgCl₂), 0.5 μ L RNasin, and 1 μ L AMV reverse transcriptase (9.500 units/ml). This was incubated at 42°C for one hour and heated to 80°C 15 for 10 min to halt the reaction. Resultant cDNA was stored at -20°C.

The preparation of primers and mutated internal standard competitive templates was as follows: ideal sequences were identified using the Oligo-TM Primer Analysis Software (National Biosciences, Hamel, Mn.). The primers were made using an Applied Biosystems 20 Model 391 PCR-Mate DNA Synthesizer. The primer sequences are described below.

Glutathione Peroxidase (GSH-Px) (Chada et al., *Genomics* 6:268-271, 1990)

The "outer" primers used to amplify both the native and mutated 25 templates result in a product length of 354 base pairs. The "outer" primers are

Sequence I.D. No. 1) (Chada et al., *Genomics* 6:268-271, 1990.)

Pos. 241 5'-GGGCCTGGTGGTGCTTCGGCT-3' (coding sense) correspond to bases 241-261 of the cloned sequence, and

30 Sequence I.D. No. 2) (Chada et al., *Genomics* 6:268-271, 1990.)

Pos. 574 5'-CAATGGTCTGGAAGCGGCGGC-3' (anti-coding sense)
which anneals to bases 574-594.

The "inner" primers used to synthesize the mutated internal standard remove an EcoRI restriction endonuclease recognition site 5 (GAATTC) by changing a native cDNA base pair (bold bases). The "inner" primers are

Sequence I.D. No. 3) (Chada et al., *Genomics* 6:268-271, 1990.)

Pos. 309 5'-ATTCT GATTTC CCTCAAGTACGTCCGGCCT-3'
(coding sense)

10 Sequence I.D. No. 4) (Chada et al., *Genomics* 6:268-271, 1990.)

Pos. 309 3'-TAAGA CTAAAG GGAGTTCATGCAGGCCGGA-5'
(anti-coding sense)

Both primers correspond to bases 309-338 of the cloned sequence. The mutation results from the substitution of a T for the 15 native A at position 316 of the sense strand. Restriction endonuclease digestion of the native GSH-Px yields products of 280 and 74 base pairs. Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) (Tso et al., *Nucleic Acids Res.* 13:2485-2502, 1985)

The "outer" primers used to amplify both the native and mutated 20 templates result in a product length of 788 or 790 base pairs. The "outer" primers are:

Sequence I.D. No. 5) (Tso et al., *Nucleic Acids Res.* 13:2485-2502, 1985.)

Pos. 46 5'-GGTCGGAGTCAACGGATTGGTCG-3' (coding sense)
25 corresponding to bases 9-32 of the cloned sequence, and

Sequence I.D. No. 6) (Tso et al., *Nucleic Acids Res.* 13:2485-2502, 1985.)

Pos. 812 5'-CCTCCGACGCCCTGCTTCACCAC-3' (anti-coding sense) which anneals bases 777-798.

The "inner" primers used to synthesize the mutated template create an EcoRI restriction endonuclease recognition site (GAATTC) by changing one native cDNA base pair (bold bases). The "inner" primers are:

5 Sequence I.D. No. 7) (Tso et al., *Nucleic Acids Res.* 13:2485-2502, 1985.)

Pos. 234 5'-TGATCAATG GAATT^C CCATCACCA-3' (coding sense)

10 Sequence I.D. No. 8) (Tso et al., *Nucleic Acids Res.* 13:2485-2502, 1985.)

Pos. 234 3'-ACTAGTTAC CTTAAG GGTAGTGGT-5' (anti-coding sense)

15 Both primers correspond to bases 199-222 of the cloned sequence. The mutation results from the substitution of a T for the native A at position 211 of the sense strand. Restriction endonuclease digestion of the mutated GAPDH yields products of 588 and 200 base pairs.

20 Several experiments were performed using a different mutated GAPDH template. This template had a novel BamHI restriction site introduced.

The "outer" primers used to amplify both the native and mutated templates result in a product length of 634 base pairs. The "outer" primers are:

25 Sequence I.D. No. 9) (Tso et al., *Nucleic Acids Res.* 13:2485-2502, 1985.)

Pos. 200 5'-CATGGCACCGTCAAGGCTGAGAAC-3' (coding sense) corresponding to bases 165-188 of the cloned sequences, and

Sequence I.D. No. 10) (Tso et al., *Nucleic Acids Res.* 13:2485-2502, 1985.)

Pos. 813 5'-CCTCCGACGCCTGCTTACCCAC-3' (anti-coding sense) which anneals to bases 777-798.

The "inner" primers used to synthesize the mutated template create a BamHI restriction endonuclease recognition site (GGATCC) by 5 changing one native cDNA base pair (bold bases). The "inner" primers are:

Sequence I.D. No. 11) (Tso et al., *Nucleic Acids Res.* 13:2485-2502, 1985.)

Pos. 368 5'-CAGGGG GGATCC AAAAGGGTCATCAT-3' (coding 10 sense)

Sequence I.D. No. 12) (Tso et al., *Nucleic Acids Res.* 13:2485-2502, 1985.)

Pos. 368 3'-GTCCCC CCTAGG TTTTCCCAGTAGTA-5' (anti-coding sense)

15 Both primers correspond to bases 333-358 of the cloned sequence. The mutation results from the substitution of a T for the native G at position 342 of the sense strand. Restriction endonuclease digestion of this mutated GAPDH yields products of 460 base pairs and 174 base pairs.

20 The mutated internal standard competitive templates were prepared by site directed mutagenesis as described by Higuchi et al., *Nucleic Acids Res.* 16:7351-7367, 1988. These single base mutations resulted in either the gain (GAPDH) or loss (GSH-Px) of an EcoRI restriction endonuclease recognition site. (Experiments were also 25 conducted using a muted GAPDH with a BamHI site introduced). For each mutated product, two initial polymerase chain reactions using an "outer" primer and an "inner" single base mismatched primer produce two overlapping DNA fragments. (Primers 1 and 4, 2 and 3 for GSH-Px; Primers 5 and 8, 6 and 7 for GAPDH). These overlapping DNA fragments 30 were electrophoresed on a 3% Nusieve, 1% LE agarose ethidium bromide

stained gel. Bands were excised and purified using Millipore Ultrafree-MC 0.45 μ M filter (Nihon Millipore Kogyo K.K., Yonezawa, Japan). The purified DNA was ethanol precipitated, washed, vacuum dried and taken up in 100 μ L sterile dH₂O. 1 μ L of each of the two overlapping DNA fragments were PCR amplified using the outer primers only. The first PCR cycle was performed without primers to allow for heterodimer formation. The entire mutated product was thus formed and amplified. The mutated PCR product was gel purified as described above and reamplified to form bulk product. The bulk product was gel purified and measured spectrophotometrically. The mutated products were diluted to the attomolar range for use as competitive templates. Herring sperm DNA (Lofstrand, Bethesda, MD) 1 μ g/ml was used as a carrier. Restriction endonuclease digestion was performed on samples of each mutated template to assure lack of contamination.

The PCR conditions were as follows: The PCR conditions were standardized for each experiment by using a master mixture containing 1x PCR buffer (50 mM KCl, 10 mM Tris-HCl, pH 9.0, 1.5 mM MgCl₂), 25 pmoles of primers coding for GSH-Px and GAPDH, 0.2 mM dNTP's (A,T,C,G), and constant amounts of both internal standards per 100 μ L reaction mixture. *Taq* DNA polymerase (2.5 units) was added to each 100 μ L reaction prior to amplification. cDNA obtained from the BEP2D cell line was serially diluted and added to the sample PCR tubes. In all experiments, control tubes containing no template, native cDNA only, or mutated standards only were amplified to check for contamination or complete enzyme digestion.

PCR amplification was carried out for 35 cycles at 94°C for one min, 60°C for one min, and 72°C for one min. After amplification, PCR products were heated for 10 min in order to maximize heterodimer formation.

The quantification of products was as follows: Samples (40 μ L) for each PCR tube were EcoRI restriction endonuclease digested for 12-16 hours (Experiments conducted using mutated GAPDH with the novel BamHI restriction site were also BamHI restriction endonuclease digested 5 for 4-5 hours). These products were isolated by electrophoresing on a 3% Nusieve, 1% LE agarose ethidium bromide stained gel for 2-3 hours at 60 V. A negative photograph was taken of the gel using Polaroid 665 positive/negative instant film.

10 The negative photograph was subjected to densitometry (Zeineh Soft Laser Scanning Densitometer Model SLR 2D/1D using Zeineh 1D Autostepover Videophoresis Program Software, Biomed Instruments, Fullerton, California). Areas under each curve were calculated and used for quantification. Corrections were made for relative band sizes and heterodimer formation. Data were expressed GSH-Px to GAPDH relative 15 ratios.

In a second set of experiments, multiplex competitive reverse transcriptase polymerase chain reaction (MC RT-PCR) with competitive templates were prepared by the Celi method to evaluate the cytochrome p450 (CYP) IAI gene in β -naphthoflavone-exposed BEP2D cells. The 20 induction of CYP1AI gene expression was evaluated using both MC RT-PCR with Celi competitive templates, and Northern analysis. Competitive templates were prepared for both the CYP1AI and GAPDH genes. The primers used to prepare the competitive template for GAPDH were: Sequence I.D. No. 13) (Tokunaga et al., *Cancer Res.* 47:5616-5619, 25 1990.)

Pos. 75 5'-GGT CGG AGT CAA CGG ATT TGG TCG-3'Pos. 94
and:

Sequence I.D. No. 14) (Tokunaga et al., *Cancer Res.* 47:5616-5619, 1990.)

Pos. 842 5'-CCT CCG ACG CCT GCT TCA CCC CAT CAC GCC
ACA GTT TCC C-3' Pos. 616

The lower outer primer used in conjunction with Sequence I.D. No. 13 to amplify both the competitive and native templates was

5 Sequence I.D. No. 15) (Tokunaga et al., *Cancer Res.* 47:5616-5619, 1990.)

Pos. 842 5'-CCT CCG ACG CCT GCT TCA CC-3' Pos. 822

The primers used to prepare the competitive template for CYPIAI were: Sequence I.D. No. 16) (Jaiswal et al., *Science* 228:80-83, 1989.)

10 Pos. 1241 5'-CAT CCC CCA CAG CAC AAC AAG-3' Pos. 1262
and:

Sequence I.D. No. 17) (Jaiswal et al., *Science* 228:80-83, 1989.)

Pos. 1555\ /Pos. 1428

Pos. 1575 5'-ACA GCA GGC ATG CTT CAT GGG TCT CAC CGA

15 TAC ACT TCC G-3' Pos. 1448

The lower outer primer used in conjunction with Sequence I.D. No. 18 to amplify both the competitive and native templates was

Sequence I.D. No. 18) (Jaiswal et al., *Science* 228:80-83, 1989.)

Pos. 1575 5'-ACA GCA GGC ATG CTT CAT GG-3' Pos. 1555

20 The PCR amplification conditions were the same as described for experiments using the competitive templates prepared for GAPDH and GSHPx by the Higuchi method except the annealing temperature was 55 degrees centigrade and the amplification was carried out for 38 cycles.

Because the native and competitive templates separate without
25 prior restriction endonuclease digestion, samples were taken directly from the PCR reaction tube and applied to ethidium bromide stained 3% Nusieve, 1% LE agarose gels. It was then possible to quantify the products by taking a negative photograph of the gel using Polaroid 665 positive/negative instant film, subjecting the negative photograph to
30 densitometry.

RNA from BEP2D cells incubated for varying time with β -naphthoflavone (10 μ M) was either electrophoresed on a 1% LE formaldehyde denaturing gel for Northern analysis or MC RT-PCR amplified, as described above. For Northern analysis, following transfer 5 of the RNA to GeneScreen, the filters were hybridized with 32 P-labeled CYPIAI cDNA.

RESULTS

The procedure used for PCR quantitation is shown schematically in Fig. 1. Serial dilutions of BEP2D cDNA (representing 0.25 μ g to 0.05 10 μ g total RNA) were co-amplified with constant amounts of each single base mutated internal standard (10 attamoles each), then analyzed as described above. A negative photograph of the gel was analyzed by densitometry in order to quantify each band, as seen in Figure 2.

Using the area under each curve, relative ratios between 15 native/mutated products were obtained. Corrections were made for relative band sizes (i.e. mutated GAPDH was multiplied by 788/588 when compared to native GAPDH and native GSH-Px was multiplied by 354/280 when compared to mutated GSH-Px). Heterodimer formation was maximized following PCR by heating the products to 100°C for 10 20 min followed by slow cooling. Following maximization of heterodimers, the quantity of each product was determined by analysis of the densitometric data using the quadratic formula as the formation of heteroduplexes follows a binomial distribution under these conditions (Gilliland et al., *Proc. Natl. Acad. Sci.* 87:2725-2729 (1990), Becker- 25 Andre et al., *Nucleic Acids Res.* 17:9437-9446 1989). Final values were expressed as an odds ratio of GSH-Px native/mutated to GAPDH native/mutated. While the graph, shown in Figure 3, of the amount of native product (in arbitrary densitometric units) vs. total starting RNA did not remain linear throughout for either GSH-Px or GAPDH, the graph of 30 the ratios of GSH-Px native/mutated vs. total starting RNA and GAPDH

native/mutated vs. total starting RNA was linear for both genes. By averaging the ratio obtained from each sample tube (2.18:1, 1.14:1, 2.00:1, 1.76:1, 2.46:1, 2.71:1, and 1.92:1), a mean value of the ratio GSH-Px native/mutated to GAPDH native/mutated of 2.17:1 with a S.D. 5 of 0.33 was obtained. No value varied more than 25% from the mean.

To assess the variability of this technique, a repeat of the above experiment was performed using different dilutions of mutated standards and master mixture. By averaging the ratio obtained from each sample tube (1:9.09, 1:8.13, 1:9.43, 1:8.13, 1:6.62, 1:8.77, 1:7.69, 1:10.00, 10 1:7.58, and 1:7.04), a mean value of the ratio of GSH-Px native/mutated to GAPDH native/mutated of 1:8.25 with a S.D. of 1.07 was obtained. No value varied more than 22% from the mean. This confirms the precision of this technique and also illustrates the variability introduced by new master mixtures containing new dilutions of mutated standards.

15 To assess the variability between samples using the same master mixture and dilutions of mutated standards, BEP2D RNA was independently extracted from three separate flasks and reverse transcribed to cDNA. Only coarse (5 fold) dilutions of cDNA were performed. Four PCR tubes were run for each study. The obtained ratios 20 of GSH-Px native/mutated to GAPDH native/mutated were 15.01:, 17.69:1, and 21.76:1 (mean=18.15, S.D.=3.40). All 3 values were within 20% of the mean. This confirms the precision of this technique when comparing samples that have been independently reverse transcribed but amplified with the same master mixture and internal 25 standard dilutions.

Northern analysis of BEP2D RNA reveals a ratio of GSHPx/GAPDH mRNA of approximately 1:8.

Example

30 Serial dilutions of BEP2D cDNA were co-amplified with constant amounts of each single base mutated internal standard competitive

templates (10 attomoles each), and then evaluated. Negative photographs of the gels (shown in Fig. 2) were analyzed by densitometry in order to quantify each band.

Starting with the area under each curve obtained by the 5 densitometric evaluation of the bands, the ratios of native/mutated products were calculated as follows. Data were evaluated from one gene at a time. Corrections were made for relative band sizes. (Mutated GAPDH was multiplied by 788/588 when compared to native GAPDH and native GSH-Px was multiplied by 354/280 when compared to 10 mutated GSH-Px). During PCR, under conditions in which primer is limiting, heterologous single strands of DNA with sequence homology may anneal to form heterodimers (Gilliland, G., Perrin, S., Blanchard, K. and Bunn, H.F. (1990) *Proc. Natl. Acad. Sci.* 87:2725-2729). When the 15 heterologous strands differ by only one base pair, as in this invention, the heterologous strands re-anneal randomly (Gilliland et al., supra; Thompson, J.D., Brodsky, I., and Yunis, J.J. (1992) *Blood* 79:1629-1635), as shown in the Punnett square below:

20

	N	M
N	NN	NM
M	NM	MM

Where N = the proportion of single-stranded native product prior to re-annealing, M = the proportion of single-stranded mutated product prior to re-annealing, NN (or N^2) = the proportion of double-stranded native product after re-annealing, 2NM = the proportion of heterodimer formed 25 after re-annealing, and MM (or M^2) = the proportion of double-stranded mutated product after re-annealing.

Heterodimers were accounted for indirectly because they were not cut by the restriction enzyme and had the same electrophoretic mobility as the undigested homodimer. Therefore, heterodimers were read 30 densitometrically along with the undigested homodimer. In order to

quantitate products, based on the Punnett square distribution, random heterodimer formation was ensured following PCR. This was done (according to the methods described in Gilliland et al., *supra*, and Thompson et al., *supra*), by heating the products to 100°C for 10 min.

5 followed by slow cooling.

For GAPDH, neither the native product (NN) nor the heterodimer (NM) were cleaved by EcoRI. Therefore, the larger band represented both native GAPDH homodimer (NN) and the NM heterodimer. This band was presented arithmetically by $N^2 + 2NM$, according to the Punnett square, while the proportion in the band resulting from EcoRI cleavage was represented by the value M^2 . Therefore, when the amount of native (N) and mutated (M) template are equal (1:1) prior to PCR, after heterodimer formation is randomized, the apparent ratio will be 3:1 [$N^2 + 2NM$]: M^2 . To illustrate this further, the raw densitometric data 10 from the first sample lane (shown in Fig. 2) are shown in Table 1 and are 15 mathematically processed to final ratios below:

The value of M^2 is known (2,214), as is the value of $N^2 + 2NM$ (10,095). From this information, M is calculated (47.05) and solving for N results in quadratic equation ($aX^2 + bX + c = 0$):

20
$$N^2 + 2N(47.05) - 10,095 = 0$$

The quadratic formula ($N = -b + /-\sqrt{b^2 - 4ac}/2a$) is used to solve for N. In this case, a = 1, b = 94.1, c = 10,095, and thus N = 63.89. The information sought is the ratio N/M which is 63.89/47.10 or 1.36/1. (Although proportions of single-stranded DNA present after PCR are 25 solved for, they are identical to those of the corresponding double-stranded DNA present prior to the PCR.)

Since densitometric values are relative, it is possible to avoid the inconvenience of using the quadratic formula by assigning the bands proportionate densitometric values that when added = 1 or 30 ($N^2 + 2NM$) + M^2 = 1. Solving for this equation:

$$(N^2 + 2NM) + M^2 = (N + M)^2 = 1 \text{ and therefore } N + M = 1$$

The relative fractions of 1 assigned to each of the bands is determined by their respective densitometric values (Table 1). Since the total densitometric value of both bands is 12,309 (10,095 + 2,214), the 5 relative proportion of the larger band ($N^2 + 2NM$) is 0.82 (10,095/12,309) and the relative proportion of the smaller band (M^2) is 0.18 (2,214/12,309). Thus, the proportion of mutated GAPDH homodimer (M^2) is 0.18, and the proportion of single-stranded mutated GAPDH (M) is 0.424. Since $N + M = 1$, the proportion of single-stranded native 10 GAPDH (N) is 1 - .424 or .576, and the ratio of native to mutated product is .576/.424 or as calculated above 1.36/1.

TABLE 1
Densitometric Data from First Sample Lane of Fig. 2

	Length (base pairs)	Zeineh densitometric values	Size correction	Relative densitometric values
(A) GAPDH				
Native GAPDH and heterodimer (N ² + 2NM)	788	10,095		0.820
Mutated GAPDH (M ²)	588	1,652	X788/588 = 2214	0.180
(B) GSH-Px				
Mutated GSH-Px and heterodimer (M ² + 2MN)	354	6,709		0.442
Native GSH-Px (N ²)	280	6,692	X354/280 = 8461	0.558

Next, the same calculations are carried out using the densitometric values for native and mutated GSH-Px from the same lane as the GAPDH values above (Table 1):

$$N^2 = .558, N = .747, \text{ and } M = 1 - .747 = .253$$

5 Native/mutated ratios are obtained:

$$\text{GSH-Px native/mutated} = .747/.253 = 2.95/1$$

$$\text{GAPDH native/mutated} = .576/.424 = 1.36/1$$

And, final odds ratios are reported:

$$\text{GSH-Px native/mutated: GAPDH native/mutated} = 2.95/1.36 = 2.17/1$$

10 As shown above, final values were expressed as an odds ratio of GSH-Px native/mutated to GAPDH native/mutated. While the relationship between the amount of native product (in arbitrary densitometric units) and total starting RNA did not remain linear throughout for either GSH-Px or GAPDH (as shown in Fig. 3), the relationship of the ratios GSH-Px
15 native/mutated and GAPDH native/mutated to total starting RNA was linear for both genes (as shown in Fig. 4). By averaging the ratio of GSH-Px native/mutated to GAPDH native/mutated obtained from sample tube (2.17:1, 2.14:1, 2.00:1, 1.76:1, 2.46:1, 2.71:1, and 1.92:1), a mean value of 2.17:1 with a S.D. of 0.33 was obtained. No value varied
20 more than 25% from the mean.

To assess the variability of this technique, the experiment was repeated using different dilutions of mutated standards and master mixture. By averaging the ratio of GSH-Px native/mutated to GAPDH native/mutated obtained from each sample tube (1:9.09, 1:8.13, 1:9.43, 25 1:8.13, 1:6.62, 1:8.77, 1:7.69, 1:10.00, 1:7.58, and 1:7.04), a mean value of 1:8.25 with a S.D. of 1.07 was obtained. No value varied more than 22% from the mean.

30 To assess the variability between samples using the same master mixture and dilutions of mutated standards (using mutated GAPDH with novel BamHI restriction site), BEP2D RNA was independently extracted

from three separate flasks and reverse transcribed to cDNA. Five fold dilutions of cDNA were performed. Four PCR tubes were run for each study. The obtained ratios of GSH-Px native-mutated to GAPDH native/mutated were 15.01:1, 17.69:1, and 21.76:1. (mean = 18.15, 5 S.D. = 3.40). All 3 values were within 20% of the mean.

The CYP1A1 gene expression was observed in both Northern and MC RT-PCR analysis to approximately the same degree, as seen in Fig. 5. By comparing the bands of the GAPDH housekeeping gene representing native and competitive template cDNA, it is clear that 10 approximately the same amount of cDNA was loaded in the lane with 1 ul of sample from control cells and the lane with 3 ul of sample from β -naphthoflavone exposed cells. Having determined which lanes were loaded the same, it is then clear that the band representing the native CYP1A1 gene is much more strongly represented in the lane containing 15 cDNA from β -naphthoflavone exposed cells compared to control cells.

Example

40 competitive templates have been combined in a single mixture, so that 40 genes can be quickly quantified (the mixture contains genes coding for xenobiotic metabolism enzymes, differentiation specific 20 proteins, transcription factors, DNA repair enzymes, apoptosis protein genes and others) in any give sample. In certain embodiments there are competitive templates for over 80 genes. In certain preferred embodiments, practitioners of this invention may not necessarily want to evaluate all of these genes simultaneously; but rather they may expect 25 to prepare mixtures of approximately 50 genes.

The greatest source of error and loss of data, and a great consumer of time, is the preparation of the PCR reactions, followed by placement in the wells for electrophoresis.

The present invention provides a method and apparatus to 30 automate the process of molecular phenotyping through quantitative RT-

OCR by making a machine with capillary conduits. The present inventive apparatus is an improvement over an oligoDNA synthesizer. According to the present invention, the PCR reactions are prepared from stocks in reservoirs connected to capillary tubing. After mixing or combining 5 appropriate aliquots of each component necessary for amplification of a particular gene from a particular cDNA, the mixture moves through capillary tubing into an air thermocycler. Then, after PCR, the mixture moves directly into wells of agarose gels. After sufficient electrophoresis, the gels are digitally imaged automatically, and the image 10 is automatically analyzed to quantify gene expression. The only manual steps needed are to program the apparatus for the desired genes to be amplified and to place the cDNA sample to be analyzed in the apparatus so that appropriate aliquots can be made. The time to amplify all selected genes is approximately 10 minutes for mixing, 30 minutes for 15 PCR amplification, 90 minutes for electrophoresis, and 10 minutes for digital imaging and data analysis, for a total time of 2 hours and 15 minutes. In certain embodiments the apparatus comprises a 40 channel device, which provides analysis of 40 genes in a little over two hours. It is also within the contemplated scope of the present invention that an 20 apparatus with 200 or 1000 or more channels can be constructed.

Fig. 6 is a schematic illustration of an apparatus of the present invention. The circuitry can comprises Y-valves 1, 2 and 4, multi-channel connector 1. The apparatus functions as follows.

Solution A in glass vial A is a stock solution of PCR reaction 25 mixture containing dNTPs, buffer, electrophoresis dye and with or without primers for β -actin or GAPDH. A capillary tube passes through an O-ring in the lid of glass vial A, then through an on/off valve and then enters Y-valve 1. The other capillary tube entering Y-valve 1 leads from multi-channel connector 1. In the embodiment shown, there are 20 30 capillary tubes leading into multi-channel connector 1. Each of the 20

capillary tubes leading into multi-channel connector 1 comes from an on/off valve. The capillary tubes leading into the on/off valves come from separate competitive template (CT) mixtures (solutions B) in glass vials after passing through O-rings in the lids. There are 4 groups of CT 5 mixtures (each group containing 100 different CTs at 5 different molarities ($10E^{-12}$ - $10E^{-16}$ M)). It is possible to change the group of CT mixtures to a different group depending on the need. The capillary tube leading from Y-valve 1 enters Y-valve 2. The other capillary tube entering Y-valve 2 comes from multi-channel connector 2. The capillary 10 tube leaving Y-valve 2 enters Y-valve 4. The other tube entering Y-valve 4 comes from multi-channel connector 3. The capillary tube leaving Y-valve 4 enters Y-valve 5.

The apparatus further comprises Y-Valve 3 and multi-channel connector 2. In the embodiment shown, there are 40 capillary tube 15 centering multi-channel connector 2 and each comes from one of 40 Y-valves 3. The capillary tube entering each Y-valve 3 comes from one of 40 glass vials each containing a solution of primers (solutions C) for a particular gene. Between the Y-valves 3 and multi-channel connector 2 are on/off valves. The other capillary leaving the Y-valves 3 enter Y- 20 valves 7.

The apparatus further comprises Y-Valves 5, 6, 7 and 8 and multi-channel connector 4. In the embodiment shown, the other capillary tube leaving each Y-valve 3 passes through an on/off valve then join Y-valves 7. The other tubes entering Y-valves 7 come from multi-channel 25 connector 4 after passing through on/off valves. The capillary entering multi-channel connector 4 comes from Y-valve 5. Ten of the tubes leaving Y-valves 7 enter Y-valves 8. The other tubes entering Y-valves 8 come from Y-valves 6 after passing through on/off valves. The tubes from the remaining 30 Y-valves 7 and the tubes leaving the 10 Y-valves 30 8 each enter a separate glass capillary tube within an air-thermocycler.

The desired solutions can be comprised of the following materials.

PCR reaction components including buffer, electrophoresis dye, glycerol, and dNTPs, and with or without actin or GAPDH primers, stored at 0°C (6 μ l for each reaction).

5 Competitive template (CT) mixtures at 10X concentration (1 μ l for each reaction). There are reservoirs for four different mixtures of 50 CTs at five different concentrations for a total of 20 CT mixture reservoirs.

10 Primers (1 μ l for each reaction). There are 50 separator reservoirs of primer solutions. Each solution contains a pair of primers at 10X concentration. The solutions come as blocks of 50 that may be changed depending on which CT solutions are planned for the experiment.

Taq polymerase solution (1 μ l for each reaction), stored at 0°C.

15 Solutions of cDNA (1 μ l used for each reaction). These can be changed for each experiment. Each cDNA solution contains a β -actin and/or GAPDH internal standard for efficiency of reverse transcription. Prior to reverse transcription, a known quantity of an RNA sequence containing a poly A tail colinear with a β -actin or GAPDH sequence is included in each RNA sample. The sequence is a shortened sequence of β -actin or GAPDH that will amplify with the primers used to amplify the 20 native sequence.

25 The use of high density oligonucleotide arrays to measure PCR Products following quantitative PT-PCR according to the methods described herein can be conducted with the following modifications. The oligonucleotide hybridizing to the sense strand of each cDNA being amplified is fluorescently labeled. One or more of the dNTP's in the PCR reaction is labeled with a fluorescent dye.

30 Preparation of high density oligonucleotide arrays can be made with the following properties. For each gene, two oligonucleotide arrays are prepared. One array has attached to it oligonucleotides (a in figure 2) that are homologous to, and will bind to, sequences unique to the

native template for a gene that was PCR-amplified using the methods described herein. The other array has attached to it oligonucleotides (b in figure 2) that are homologous to and will bind to sequences that span the juncture between the 5' end of the competitive template, and the 5 truncated, mis-aligned 3' end of the competitive template that was prepared according to the method of Celi.

The PCR products are applied to the high density oligonucleotide arrays prepared according to the method described above. The expression of the target genes is quantified by comparing the fluorescent 10 intensities of the arrays for the native and CT for the housekeeping genes and targets genes.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Willey, James C
- (ii) TITLE OF INVENTION: Method for Quantitative Measurement of Gene Expression Using Multiplex Competitive Reverse Transcriptase-Polymerase Chain Reaction
- (iii) NUMBER OF SEQUENCES: 18
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Emch, Schaffer, Schaub, & Porcello Co., L.P.A.
 - (B) STREET: One Seagate, Suite 1980
 - (C) CITY: Toledo
 - (D) STATE: Ohio
 - (E) COUNTRY: USA
 - (F) ZIP: 43604
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/188,434
 - (B) FILING DATE: 28-Jan-1994
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/043,390
 - (B) FILING DATE: 06-Apr-1993
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Martineau, Catherine B.
 - (B) REGISTRATION NUMBER: 31,854
 - (C) REFERENCE/DOCKET NUMBER: 93206
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (419) 243-1294
 - (B) TELEFAX: (419) 243-8502

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO

38

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo Sapiens
- (H) CELL LINE: HL-60

(x) PUBLICATION INFORMATION:

- (A) AUTHORS: Chada, S.
Le Beau, M. M.
Casey, L.
Newberger, P. E.
- (C) JOURNAL: Genomics
- (D) VOLUME: 6
- (F) PAGES: 268-271
- (G) DATE: September-1990

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGGGCCTGGT GGTGCTCGGC T

21

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo Sapiens
- (H) CELL LINE: HL-60

(x) PUBLICATION INFORMATION:

- (A) AUTHORS: Chada, S.
Le Beau, M. M.
Casey, L.
Newberger, P. E.
- (C) JOURNAL: Genomics
- (D) VOLUME: 6
- (F) PAGES: 268-271
- (G) DATE: September-1990

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CAATGGTCTG GAAGCGGCAG C

21

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

39

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo Sapiens
- (D) DEVELOPMENTAL STAGE: Adult
- (H) CELL LINE: HL-60

(x) PUBLICATION INFORMATION:

- (A) AUTHORS: Chada, S.
Le Beau, M. M.
Casey, L.
Newberger, P. E.
- (C) JOURNAL: Genomics
- (D) VOLUME: 6
- (F) PAGES: 268-271
- (G) DATE: 1990

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATTCTGATT CCCTCAAGTA CGTCCGGCCT

30

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo Sapiens
- (D) DEVELOPMENTAL STAGE: Adult
- (H) CELL LINE: HL-60

(x) PUBLICATION INFORMATION:

- (A) AUTHORS: Chada, S.
Le Beau, M. M.
Casey, L.
Newberger, P. E.
- (C) JOURNAL: Genomics
- (D) VOLUME: 6

40

(F) PAGES: 268-271
(G) DATE: 1990

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TAAGACTAAA GGGAGTTCAT GCAGGCCGGA

30

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Adult
(F) TISSUE TYPE: Liver
(G) CELL TYPE: Hepatocellular

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Tso, J. Y.
 Sun, X.
 Kao, T.
 Reese, K. S.
 Wu, R.
(C) JOURNAL: Nucleic Acids Research
(D) VOLUME: 13
(F) PAGES: 2485-2502
(G) DATE: 1985

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGTGGGGAGT CAACGGATTG GGTGCG

25

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

41

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Adult
(F) TISSUE TYPE: Liver
(G) CELL TYPE: Hepatocellular

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Tso, J. Y.
Sun, X.
Kao, T.
Reece, K. S.
Wu, R.
(C) JOURNAL: Nucleic Acids Research
(D) VOLUME: 13
(F) PAGES: 2485-2502
(G) DATE: 1985

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCTCCGACGC CTGCTTCACC AC

22

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Adult
(F) TISSUE TYPE: Liver
(G) CELL TYPE: Hepatocellular

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Tso, J. Y.
Sun, X.
Kao, T.
Reece, K. S.
Wu, R.
(C) JOURNAL: Nucleic Acids Research
(D) VOLUME: 13
(F) PAGES: 2485-2502
(G) DATE: 1985

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TGATCAATGG AATTCCCATC ACCA

24

(2) INFORMATION FOR SEQ ID NO:8:

42

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo Sapiens
 - (D) DEVELOPMENTAL STAGE: Adult
 - (F) TISSUE TYPE: Liver
 - (G) CELL TYPE: Hepatocellular
- (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Tso, J. Y.
 - Sun, X.
 - Kao, T.
 - Reece, K. S.
 - Wu, R.
 - (C) JOURNAL: Nucleic Acids Research
 - (D) VOLUME: 13
 - (F) PAGES: 2485-2502
 - (G) DATE: 1985

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ACTAGTTACC TTAAGGGTAG TGGT

24

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo Sapiens
 - (D) DEVELOPMENTAL STAGE: Adult
 - (F) TISSUE TYPE: Liver
 - (G) CELL TYPE: Hepatocellular
- (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Tso, J. Y.
 - Sun, X.

Kao, T. 43
Reece, K. S.
Wu, R.

(C) JOURNAL: Nucleic Acids Research
(D) VOLUME: 13
(F) PAGES: 2485-2502
(G) DATE: 1985

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CATGGCACCG TCAAGGCAAC

20

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Adult
(F) TISSUE TYPE: Liver
(G) CELL TYPE: Hepatocellular

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Tso, J. Y.
Sun, X.
Kao, T.
Reece, K. S.
Wu, R.

(C) JOURNAL: Nucleic Acids Research
(D) VOLUME: 13
(F) PAGES: 2485-2502
(G) DATE: 1985

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCTCCGACGC CTGCTTCACC AC

22

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

44

- (iv) ANTI-SENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo Sapiens
 - (D) DEVELOPMENTAL STAGE: Adult
 - (F) TISSUE TYPE: Liver
- (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Tso, J. Y.
Sun, X.
Kao, T.
Reece, K. S.
Wu, R.
 - (C) JOURNAL: Nucleic Acids Research
 - (D) VOLUME: 13
 - (F) PAGES: 2485-2502
 - (G) DATE: 1985

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CAGGGGGGAT CCAAAAGGGT CATCAT

26

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo Sapiens
 - (D) DEVELOPMENTAL STAGE: Adult
 - (F) TISSUE TYPE: Liver
- (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Tso, J. Y.
Sun, X.
Kao, T.
Reese, K. S.
Wu, R.
 - (C) JOURNAL: Nucleic Acids Research
 - (D) VOLUME: 13
 - (F) PAGES: 2485-2502
 - (G) DATE: 1985

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTCCCCCCTA GGTTTCCCA GTAGTA

26

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(F) TISSUE TYPE: Liver

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Tokunaga, K.
Nakamura, Y.
Sakata, K.
Fujimori, K.
Ohkubo, M.
Sawada, K.
Sakiyama, S.

(B) TITLE: Enhanced expression of a
glyceraldehyde-3-phosphate dehydrogenase gene in
human lung cancers

(C) JOURNAL: Cancer Res.
(D) VOLUME: 47
(F) PAGES: 5616-5619
(G) DATE: SUMMER-1990

(K) RELEVANT RESIDUES IN SEQ ID NO:13: FROM 1 TO 24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGTCGGAGTC AACGGATTTG GTCG

24

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Adult
(F) TISSUE TYPE: Lung
(G) CELL TYPE: Bronchial epithelial

(H) CELL LINE: Lung cancer **46**

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Tokunaga, K.
Nakamura, Y.
Sakata, K.
Fujimora, K.
Ohkubo, M.
Sawada, K.
Sakiyama, S.

(B) TITLE: Enhanced expression of a
glyceraldehyde-3-phosphate dehydrogenase gene in
human lung cancers

(C) JOURNAL: Cancer Res.

(D) VOLUME: 47

(F) PAGES: 5616-5619

(G) DATE: Summer-1990

(K) RELEVANT RESIDUES IN SEQ ID NO:14: FROM 1 TO 40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CCTCCGACGC CTGCTTCACC CCATCACGCC ACAGTTCCCC

40

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Adult
(F) TISSUE TYPE: Lung
(G) CELL TYPE: Bronchial epithelial
(H) CELL LINE: Lung cancer

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Tokunaga, K.
Nakamura, Y.
Sakata, K.
Fujimora, K.
Ohkubo, M.
Sawada, K.
Sakiyama, S.

(B) TITLE: Enhanced expression of a
glyceraldehyde-3-phosphate dehydrogenase gene in
human lung cancers

(C) JOURNAL: Cancer Res.

(D) VOLUME: 47

(F) PAGES: 5616-5619

(G) DATE: summer-1990

47

(K) RELEVANT RESIDUES IN SEQ ID NO:15: FROM 1 TO 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CCTCCGACGC CTGCTTCACC

20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo Sapiens
- (D) DEVELOPMENTAL STAGE: Adult
- (F) TISSUE TYPE: Breast
- (G) CELL TYPE: Mammary epithelial
- (H) CELL LINE: MCF-7

(x) PUBLICATION INFORMATION:

- (A) AUTHORS: Jaiswal, A. K.
Gonzalez, F. J.
Nebert, D. W.
- (B) TITLE: Human dioxin-inducible cytochrome P-1-450
(TCDD-inducible) mRNA, complete cds.
- (C) JOURNAL: Science
- (D) VOLUME: 228
- (F) PAGES: 80-83
- (G) DATE: 15-JUN-1989

(K) RELEVANT RESIDUES IN SEQ ID NO:16: FROM 1 TO 21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CATCCCCCAC AGCACAAACAA G

21

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

48

- (A) ORGANISM: Homo Sapiens
- (D) DEVELOPMENTAL STAGE: Adult
- (F) TISSUE TYPE: Breast
- (G) CELL TYPE: Mammary epithelial
- (H) CELL LINE: MCF-7

(x) PUBLICATION INFORMATION:

- (A) AUTHORS: Jaiswal, A. K.
Gonzalez, F. J.
Nebert, D. W.
- (B) TITLE: Human dioxin-inducible cytochrome P1-450:
Complementary DNA and amino acid sequence
- (C) JOURNAL: Science
- (D) VOLUME: 228
- (F) PAGES: 80-83
- (G) DATE: 15-JUN-1989
- (K) RELEVANT RESIDUES IN SEQ ID NO:17: FROM 1 TO 40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ACAGCAGGCA TGCTTCATGG GTCTCACCGA TACACTTCCG

40

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo Sapiens
- (D) DEVELOPMENTAL STAGE: Adult
- (F) TISSUE TYPE: Breast
- (G) CELL TYPE: Mammary epithelial
- (H) CELL LINE: MCF-7

(x) PUBLICATION INFORMATION:

- (A) AUTHORS: Jaiswal, A. K.
Gonzalez, F. J.
Nebert, D. W.
- (B) TITLE: Human dioxin-inducible cytochrome P1-450:
Complementary DNA and amino acid sequence
- (C) JOURNAL: Cancer Res.
- (D) VOLUME: 228
- (F) PAGES: 80-83
- (G) DATE: 15-JUN-1989
- (K) RELEVANT RESIDUES IN SEQ ID NO:18: FROM 1 TO 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ACAGCAGGCA TGCTTCATGG

20

5

WE CLAIM:

1. A method for quantitative measurement of gene expression of a target gene comprising:
 - a) isolating cellular mRNA of target genes and 10 housekeeping genes which are reverse transcribed and specifically amplified in the presence of competitive templates such that a ratio of each target gene to each housekeeping gene is obtained and that ratio is used to assess the amount of gene expression of each target gene by conducting simultaneous polymerase chain reaction amplification of a 15 mixture of the following:
 - i) at least one oligonucleotide primer pair of each target gene,
 - ii) at least one oligonucleotide primer pair of each housekeeping gene,
 - 20 iii) at least one mutated competitive template of each target gene,
 - iv) at least one mutated competitive template of each housekeeping gene, and
 - v) native cDNA which contains at least one copy 25 of each target gene cDNA and at least one copy of each housekeeping gene cDNA;
- to form polymerase chain reaction cDNA products comprising: native cDNA of each target gene and each housekeeping gene, and mutated cDNA of each target gene and each housekeeping gene; in which an 30 appropriate amount of a mixture of competitive templates for different

target genes at known concentrations relative to one another and that contain different quantified amounts of housekeeping genes relative to target genes is included in the PCR amplification mixture such that the native and competitive templates of the target genes and housekeeping genes being assessed can both be visualized following amplification by i) preparing a sufficient master mixture containing aliquots of cDNA, competitive template mixture, dNTPs, thermostable DNA polymerase and buffer for the number of genes to be measured; ii) placing an aliquot of the reaction mixture in a number of reaction vessels on ice corresponding to the number of genes to be evaluated, and iii) placing an aliquot of primers specific to one of the genes to be evaluated in each reaction vessel and placing the primers for the housekeeping gene in a separate tube,

(b) isolating the cDNA products;

(c) detecting the relative presence of the native cDNA products and mutated cDNA products by comparing the amount of native cDNA coding for each target gene and the amount of mutated cDNA coding for the competitive template of target gene to the amount of native cDNA coding for each housekeeping gene and the amount of mutated cDNA coding for the competitive template of the housekeeping gene; wherein comparison of the relative presence of the amounts of native cDNA products and mutated cDNA products of the target gene and the housekeeping gene provide the quantitative measurement of the expression of the target gene; and,

(d) measuring the ratio of the density of the band corresponding to the PCR product for the native gene relative to the density of the band corresponding to the PCR product for the competitive template for each gene; wherein, for each gene, the native/competitive template density ratio for the target gene is divided by the

native/competitive template density ratio for the housekeeping gene to provide a final value in units of mRNAs/10⁶ β -actin mRNAs.

2. The method of claim 1, in which the mutated competitive
5 templates of the target gene and of the housekeeping gene: a) comprise at least one oligonucleotide that contains a sequence homologous to the target gene or the housekeeping gene, respectively; and b) contain at least one known restriction endonuclease recognition site sequence or a point mutation for the restriction endonuclease recognition sequence.

10

3. The method of claim 2, in which the point mutation comprises one or two base pair mismatch.

4. The method of claim 3, in which the point mutation causes
15 either a gain or a loss of an EcoRI restriction endonuclease recognition site.

5. The method of claim 1, in which the target gene and the housekeeping gene have the same recognition site.

20

6. The method of claim 1, in which the cDNA products are isolated by subjecting the cDNA products to digestion with a restriction enzyme, and thereafter subjecting the digested cDNA products to electrophoresis to separate the native cDNA products from the mutated
25 cDNA products.

7. The method of claim 6, in which the mutated competitive template of the target gene causes a loss of an EcoRI restriction site and the mutated competitive template of the housekeeping gene causes a
30 gain of an EcoRI restriction site.

8. The method of claim 7, in which the cDNA products are digested with EcoRI.

9. The method of claim 1, in which the mutated competitive 5 templates of the target gene and the housekeeping gene comprise artificially shortened competitive templates.

10. The method of claim 1, in which the polymerase chain reaction amplification is allowed to proceed to a point which maximizes 10 heterodimer formation.

11. The method of claim 1, in which at least one housekeeping gene is selected from the group comprising glyceraldehyde phosphate dehydrogenase (GAPDH), β -actin or 28S RNA.

15 12. The method of claim 1, in which the mutated internal standard competitive templates are prepared by site directed mutagenesis.

20 13. The method of claim 1, in which prior to the simultaneous polymerase chain reaction amplification for each mutated competitive template:

- 25 i) two initial polymerase chain reactions are conducted using an outer primer and an inner single base mismatched mutated internal standard competitive template primer to produce two overlapping DNA fragments;
- ii) isolating and purifying the overlapping DNA fragments of the initial polymerase chain reaction of step (i) above;
- iii) polymerase chain reaction amplifying each of the two 30 overlapping DNA fragments using the outer primers only;

- iv) conducting a polymerase chain reaction amplification according to claim 7 above without the primers to allow for heterodimer formation; and
- 5 v) purifying and amplifying the polymerase chain reaction products of step (iv) above and thereafter diluting to use as competitive templates.

14. A method for quantitative measurement of gene expression of different target genes in tissue samples comprising:

- 10 a) synthesizing at least one oligonucleotide primer pair of at least one housekeeping gene,
- b) synthesizing at least one oligonucleotide primer pair of each target gene,
- c) synthesizing at least one competitive template of each 15 housekeeping gene,
- d) synthesizing at least one competitive template of each target gene,
- e) isolating at least a portion of an RNA sequence from said tissue samples,
- 20 f) subjecting the RNA sequence to reverse transcription to obtain at least one native cDNA,
- g) conducting polymerase chain reaction amplification of the native cDNA in the presence of each of the target gene oligonucleotide primers, each of the housekeeping gene oligonucleotide 25 primers, and predetermined quantities of each housekeeping gene competitive template and each target gene competitive template, in which an appropriate amount of a mixture of competitive templates for different target genes at known concentrations relative to one another and that contain different quantified amounts of housekeeping genes 30 relative to target genes is included in the PCR amplification mixture such

that the native and competitive templates of the target genes and housekeeping genes being assessed can both be visualize following amplification by i) preparing a sufficient master mixture containing aliquots of cDNA, competitive template mixture, dNTPs, thermostable 5 DNA polymerase and buffer for the number of genes to be measured; ii) placing an aliquot of the reaction mixture in a number of reaction vessels on ice corresponding to the number of genes to be evaluated, and iii) placing an aliquot of primers specific to one of the genes to be evaluated in each reaction vessel and placing the primers for the housekeeping 10 gene in a separate tube,

h) subjecting amplified cDNA products of step "g" above to digestion with at least one restriction enzyme,

i) subjecting the digested cDNA to electrophoresis to separate the amplified target gene and the amplified housekeeping gene 15 from the amplified target gene competitive template and the amplified housekeeping gene competitive template, and

j) measuring the relative expression of the target gene in at least one tissue sample by dividing a ratio of the target gene to a known amount of the competitive template of the target gene by a ratio 20 of housekeeping gene to a known amount of the competitive template of the housekeeping gene to provide the quantitative measurement of expression of the target gene by measuring the ratio of the density of the band corresponding to the PCR product for the native gene relative to the density of the band corresponding to the PCR product for the competitive 25 template for each gene; wherein, for each gene, the native/competitive template density ratio for the target gene is divided by the native/competitive template density ratio for the housekeeping gene to provide a final value in units of mRNAs/ 10^6 β -actin mRNAs.

15. The method of claim 14, which further comprises making artificially shortened competitive templates of the housekeeping gene to serve as an internal mutated standard competitive template of the housekeeping gene such that there is no need for subjecting the amplified
5 cDNA to restriction enzyme digestion.

16. The method of claim 14, in which polymerase chain reaction products are denatured at about 94°C and slowly cooled to allow heterodimer formation.
10

17. An apparatus for quantitatively measuring gene expression of target genes of interest comprising a means which is programmed to automatically mix a master mixture and competitive template mixtures such that the desired set of genes can be analyzed on any particular
15 cDNA sample.

18. The apparatus of claim 17, which comprises a means for molecular phenotyping through quantitative RT-OCR comprising a means for preparing PCR reactions from stocks in reservoirs connected to
20 capillary tubing;

a means for combining appropriate aliquots of each component necessary for amplification of a particular gene from a particular cDNA to form a mixture;

25 a means for moving the mixture through the capillary tubing into an air thermocycler,

a means for moving the mixture, after PCR, directly into wells of agarose gels;

a means to digitally image the gels; and,

30 a means to analyze the digital images to quantify gene expression.

19. The apparatus of claim 18 wherein the apparatus comprises a circuitry including

- i) Y-valves 1, 2 and 4 and multi-channel connector 1 wherein a solution A in glass vial A is a stock solution of PCR reaction mixture containing dNTPs, buffer, electrophoresis dye and with or without primers for β -actin or GAPDH, and a capillary tube passes through an O-ring in a lid of glass vial A, through an on/off valve and enters Y-valve 1; another capillary tube entering Y-valve 1 leads from multi-channel connector 1;
- 10 wherein multiple capillary tubes lead into multi-channel connector 1, each of the capillary tubes leading into multi-channel connector 1 come from an on/off valve, the capillary tubes leading into the on/off valves coming from separate competitive template (CT) mixtures (solutions B) in glass vials after passing through O-rings in the lids, wherein there are 4 groups of CT mixtures (each group containing 100 different CTs at 5 different molarities ($10E^{-12}$ - $10E^{-16}$ M)), and a means for changing the group of CT mixtures to a different group; wherein the capillary tube leading from Y-valve 1 enters Y-valve 2, the other capillary tube entering Y-valve 2 comes from multi-channel connector 2, the capillary tube leaving Y-valve 2 enters Y-valve 4, the other tube entering Y-valve 4 comes from multi-channel connector 3, and the capillary tube leaving Y-valve 4 enters Y-valve 5.
- 20 ii) Y-Valve 3 and multi-channel connector 2, wherein multiple capillary tube enter multi-channel connector 2 and each comes from one of Y-valves 3, the capillary tube entering each Y-valve 3 comes from one of glass vials each containing a solution of primers for a particular gene, wherein between the Y-valves 3 and multi-channel connector 2 are on/off valves, and the other capillary leaving the Y-valves 3 enters Y-valves 7; and

iii) Y-Valves 5, 6, 7 and 8 and multi-channel connector 4 wherein the other capillary tube leaving each Y-valve 3 passes through an on/off valve then joins Y-valves 7, the other tubes entering Y-valves 7 come from multi-channel connector 4 after passing through on/off valves, the capillary entering multi-channel connector 4 comes from Y-valve 5, one-fourth of the tubes leaving Y-valves 7 enter Y-valves 8, the other tubes entering Y-valves 8 come from Y-valves 6 after passing through on/off valves, the tubes from the remaining three-fourths of the Y-valves 7 and the tubes leaving the Y-valves 8 each enter a separate glass capillary tube within the air-thermocycler.

20. The apparatus of claim 19 wherein the solution A comprises PCR reaction components including buffer, electrophoresis dye, glycerol, and dNTPs, and with or without actin or GAPDH primers, stored at 0°C
15 (6µl for each reaction).

21. The apparatus of claim 19, wherein the solution B comprises competitive template (CT) mixtures at 10X concentration (1 µl for each reaction), and further including reservoirs for four different mixtures of
20 50 CTs at five different concentrations for a total of 20 CT mixture reservoirs.

22. The apparatus of claim 19, wherein the solution C comprises primers (1 µl for each reaction), and further including 50 separator
25 reservoirs of primer solutions wherein each solution contains a pair of primers at 10X concentration.

23. The apparatus of claim 19, which uses a Taq polymerase solution (1 µl for each reaction), stored at 0°C.

24. The apparatus of claim 19, which uses solutions of cDNA (1 μ l used for each reaction) that are changed for each experiment wherein each cDNA solution contains a β -actin and/or GAPDH internal standard for efficiency of reverse transcription and wherein prior to reverse transcription, a known quantity of an RNA sequence containing a poly A tail colinear with a β -actin or GAPDH sequence is included in each RNA sample, and wherein the sequence is a shortened sequence of β -actin or GAPDH that amplifies with the primers used to amplify the native sequence.

10

25. The apparatus of claim 19, which further includes high density oligonucleotide arrays to measure PCR products following quantitative PT-PCR.

15

26. The apparatus of claim 25, in which the oligonucleotide hybridizing to the sense strand of each cDNA being amplified is fluorescently labeled.

20

27. The apparatus of claim 25, in which one or more of the dNTP's in the PCR reaction is labeled with a fluorescent dye.

25

28. The apparatus of claim 25, wherein the high density oligonucleotide arrays have the following properties:

for each gene, two oligonucleotide arrays are prepared;

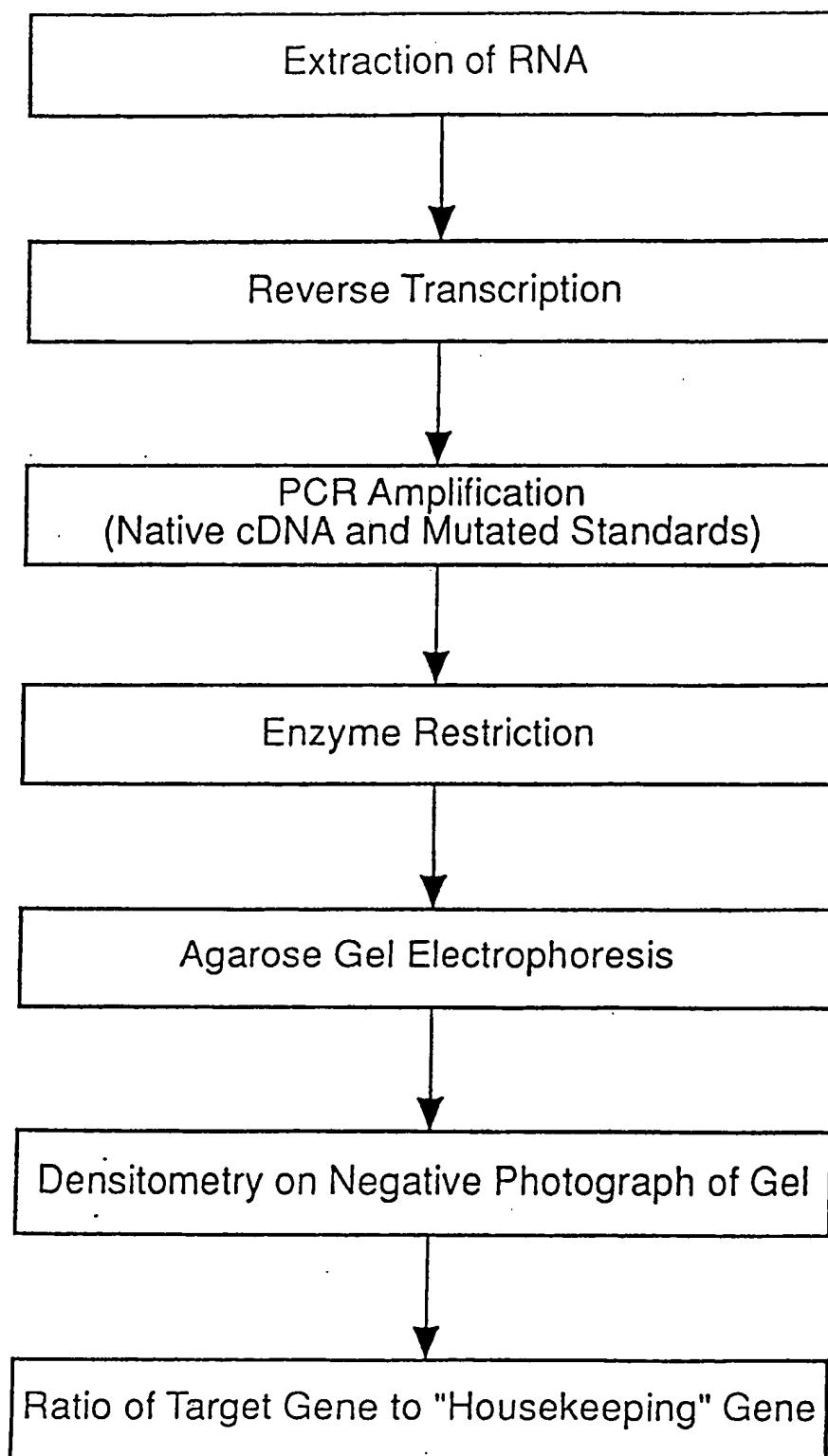
i) one array has attached to it oligonucleotides that are homologous to, and will bind to, sequences unique to the native template for a gene that was PCR-amplified; and

ii) another array has attached to it oligonucleotides that are homologous to and will bind to sequences that

span the juncture between the 5' end of the competitive template, and the truncated, mis-aligned 3' end of the competitive template.

29. The apparatus of claim 28, in which the PCR products are
5 applied to the high density oligonucleotide arrays.

30. The apparatus of claim 29, in which the expression of the target genes are quantified by comparing the fluorescent intensities of the arrays for the native and CT for the housekeeping genes and targets
10 genes.



— FIG. 1

+

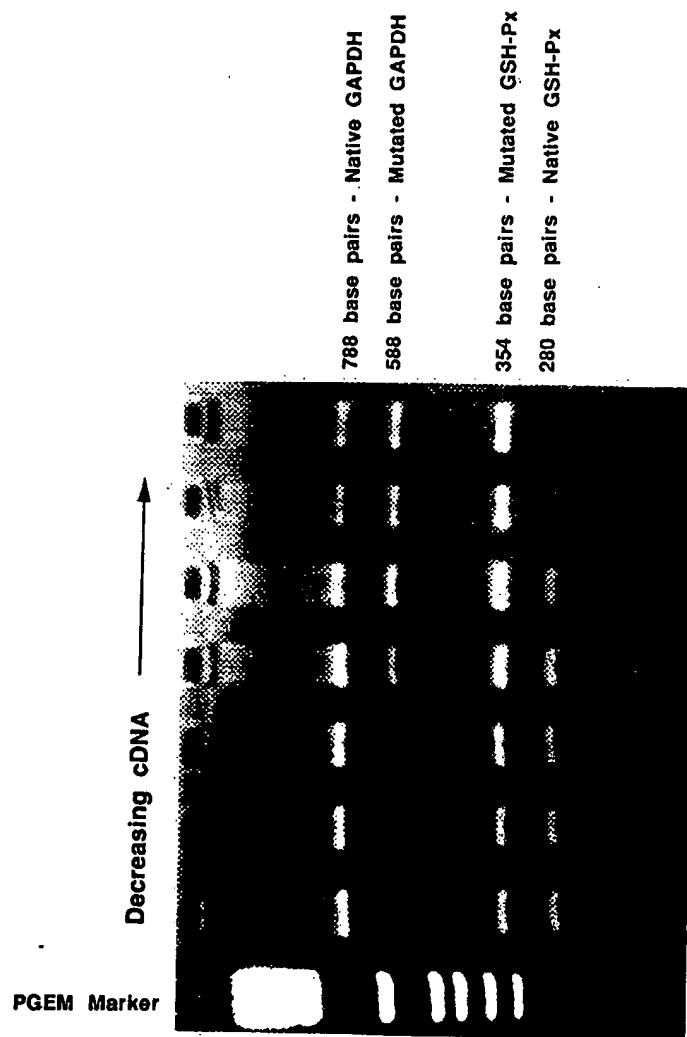


FIG. 2

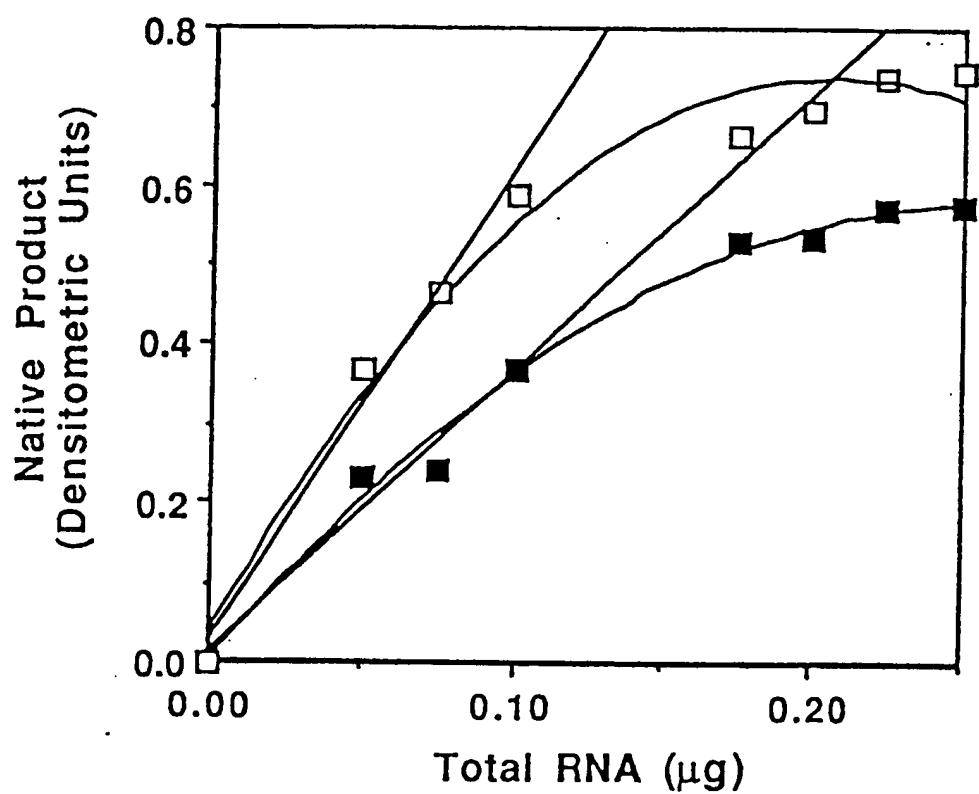


FIG. 3

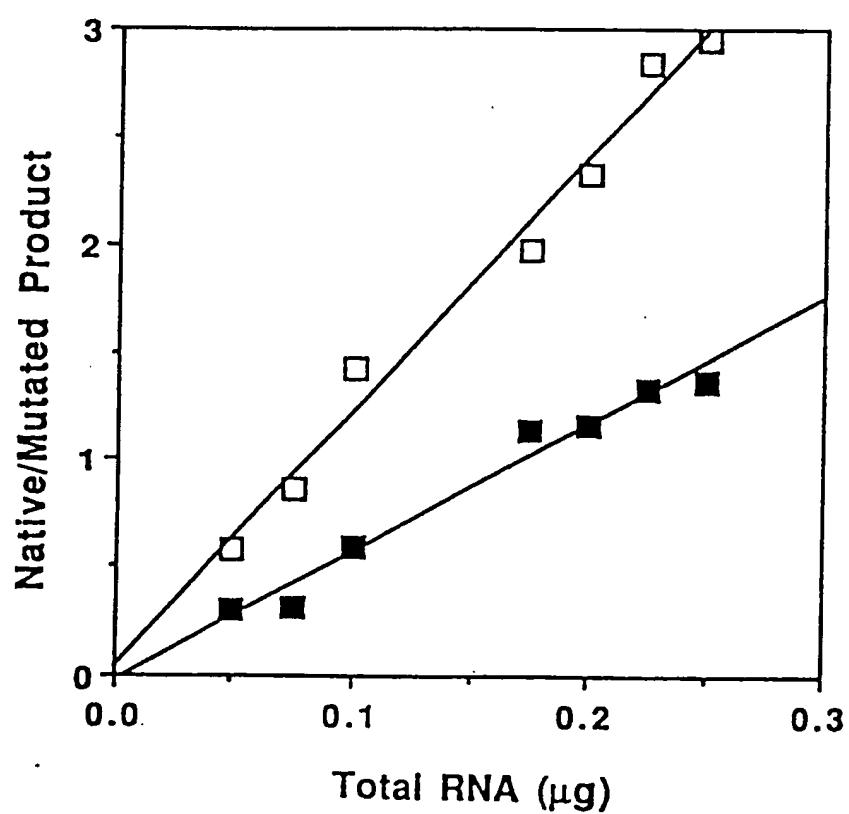
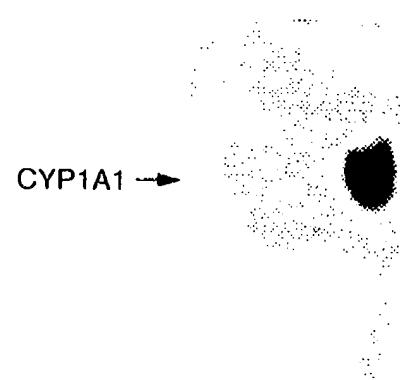


FIG. 4

5/7

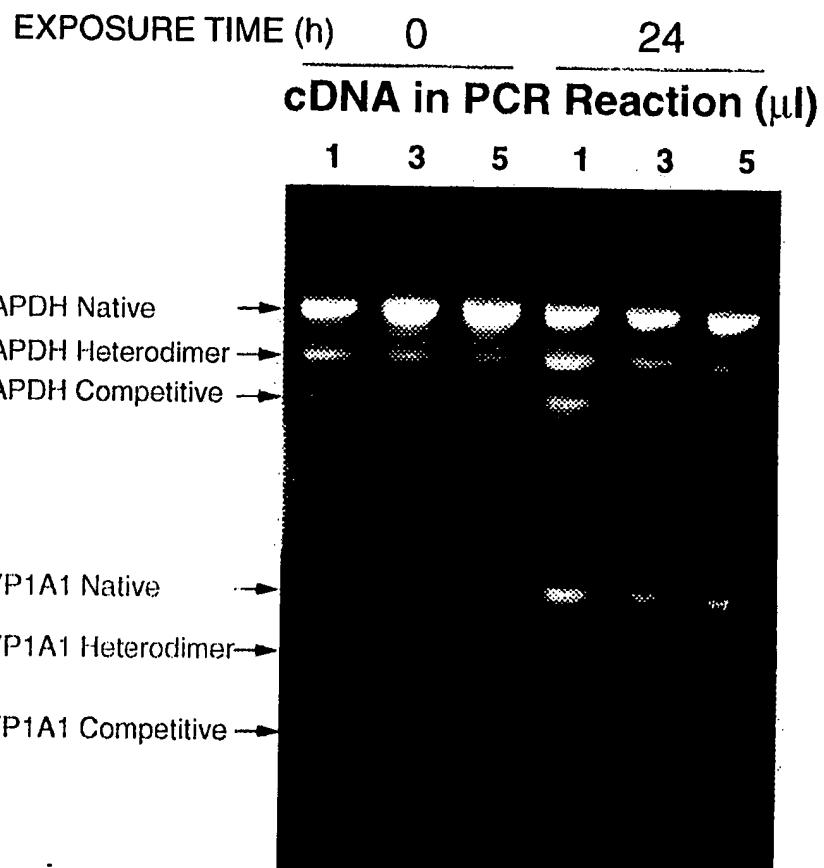
NORTHERN ANALYSIS

0 8 24 EXPOSURE TIME



—FIG. 5A

6/7

MULTIPLEX RT-PCR ANALYSIS**FIG. 5B**

